

VU Research Portal

Measuring functional limitations in rising and sitting down: Development of a questionnaire

Roorda, L.D.; Roebroeck, M.E.; Lankhorst, G.J.; van Tilburg, T.G.; Bouter, L.M.

published in

Archives of Physical Medicine and Rehabilitation
1996

DOI (link to publisher)

[10.1016/S0003-9993\(96\)90005-2](https://doi.org/10.1016/S0003-9993(96)90005-2)

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Roorda, L. D., Roebroeck, M. E., Lankhorst, G. J., van Tilburg, T. G., & Bouter, L. M. (1996). Measuring functional limitations in rising and sitting down: Development of a questionnaire. *Archives of Physical Medicine and Rehabilitation*, 77, 663-669. [https://doi.org/10.1016/S0003-9993\(96\)90005-2](https://doi.org/10.1016/S0003-9993(96)90005-2)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Measuring Functional Limitations in Rising and Sitting Down: Development of a Questionnaire

Leo D. Roorda, MD, PT, Marij E. Roebroek, PhD, Gustaaf J. Lankhorst, MD, PhD, Theo van Tilburg, PhD, Lex M. Bouter, PhD

ABSTRACT. Roorda LD, Roebroek ME, Lankhorst GJ, van Tilburg T, Bouter LM. Measuring functional limitations in rising and sitting down: development of a questionnaire. *Arch Phys Med Rehabil* 1996;77:663-9.

Objective: Develop and test a self-administered questionnaire that measures perceived and actual functional limitations in rising and sitting down.

Setting: Private practices for physical therapy and outpatient clinics of hospitals and rehabilitation centers.

Patients: 345 outpatients (43% male, aged 14 to 92 years) with different grades of functional limitations and different types of lower extremity orthopedic or rheumatologic disorders.

Methods: The Questionnaire Rising and Sitting Down (QR&S) was developed on the basis of a literature review and careful operationalization of functional limitations. Five dimensions concerning different objects (high chair, low chair, toilet, bed, and car) and one global dimension were postulated to be contained in the instrument. Mokken scale analysis was used to test the postulated dimensions (scalability coefficient H). Furthermore, robustness with respect to patient characteristics was determined, as well as intratester reliability (reliability coefficient Rho), test-retest reliability (intraclass correlation coefficient [ICC]), content validity (coverage of operationalized aspects), and construct validity (testing of seven hypotheses).

Results: Mokken scale analysis confirmed the existence of 5 object dimensions ($H = .53-.59$). However, two global dimensions were found ($H = .50-.54$). The resulting hierarchical scales, consisting of subsets of the 32 final QR&S items, are robust and measure functional limitations in a reliable (Rho .77-.91; ICC .72-.90) and valid (3 out of 4 aspects covered, 2 hypotheses rejected for 3 out of 7 scales) manner.

Conclusion: The QR&S is a reliable and valid self-administered questionnaire. It consists of hierarchical scales and measures perceived and actual functional limitations in rising and sitting down.

© 1996 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

RISING AND SITTING DOWN are basic mobility skills¹ that play a critical role in independence and mobility. Problems with rising and sitting down are especially prevalent in

the arthritic and the elderly. Munton and coworkers² found that 42% of a study population in this category reported difficulty with rising from and 23% with sitting down into their easy chairs. In the noninstitutionalized Dutch population, 7% of the general population and 36% of the elderly population (those 85 years and older) reported difficulty with rising or sitting down.³

Although rising and sitting down is of critical importance in maintaining independence, there has been little insight into the actual functional limitations that patients perceive at home when rising and sitting down. Furthermore, little is known about factors that contribute to the problems in rising and sitting down and the prevalence of different grades of limitations in various patient groups. It appears, then, that an assessment instrument is needed to study these phenomena adequately.

To our knowledge, such a comprehensive and detailed instrument is not yet available. Existing instruments⁴⁻⁷ only provide a global measure of functional limitations in rising and sitting down, because these limitations are not the primary goal of measurement. Other instruments⁸⁻¹¹ are disease-specific, making valid comparisons of results across diagnostic categories impossible. Some instruments^{1,12-17} measure patients' performance competence in a laboratory setting; however, competence in a laboratory setting does not necessarily indicate competent performance in a patient's own environment¹⁸ and the consequent absence of problems at home. Similarly, incompetence in a laboratory setting does not necessarily indicate perceived problems in the patients' own environment.

Since existing instruments did not meet our needs, we decided to develop a questionnaire to assess perceived and actual functional limitations in rising and sitting down at home. The questionnaire was intended to be a discriminative index,¹⁹ measuring cross-sectional differences between patients, to distinguish between patients with varying degrees of limitations in rising and sitting down. Furthermore, the questionnaire had to be applicable in clinical and survey studies. For clinical studies the questionnaire had to provide a comprehensive and detailed picture of limitations, while for survey studies a more global picture would suffice. Finally, the questionnaire had to be applicable to outpatients with different degrees of functional limitations and different types of orthopedic or rheumatological lower extremity disorders in order to make comparisons across diagnostic categories.

This article reports on the development and testing of the Questionnaire Rising & Sitting Down (QR&S). The testing involved the study of scalability, robustness, and reliability, as well as content and construct validity.

METHODS

Item Selection

Items were derived from a literature review on rising and sitting down. On the basis of this review, and to create a comprehensive and detailed measure of limitations in rising and sitting down, it was decided to measure limitations in rising and sitting down regarding five objects—(1) high chair, (2) low chair, (3) toilet, (4) bed, and (5) car—thereby assuming 5 object

From the Department of Rehabilitation Medicine, University Hospital Vrije Universiteit (Drs. Roorda, Roebroek, Lankhorst), and the Department of Social Research Methodology (Dr. van Tilburg) and Department of Epidemiology and Biostatistics (Dr. Bouter), Vrije Universiteit, Amsterdam, The Netherlands.

Submitted for publication July 14, 1995. Accepted in revised form December 27, 1995.

Supported in part by a grant from the St. Jorisstichting.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit upon the authors or upon any organization with which the authors are associated.

Reprint requests to Leo D. Roorda, MD, PT, Department of Rehabilitation Medicine, University Hospital Vrije Universiteit, PO Box 7057, 1007 MB Amsterdam, The Netherlands.

© 1996 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation
0003-9993/96/7707-3593\$3.00/0

dimensions. Furthermore, one global dimension was postulated to be present which underlies the five object dimensions. For each of the five objects, the aspects (1) velocity, (2) use of arm(rest)s, (3) use of help, and (4) "other differences" in performance (eg, shifting forward before rising) were operationalized. Items were formulated in behavioral terms. Dichotomous response options were chosen to facilitate uniform interpretation.¹⁹

A draft version of the QR&S was pretested twice and subjected to the opinions of experts (physicians, physical and occupational therapists, and sociologists) resulting in rewording of some items and patient instruction. This revision resulted in a questionnaire with 54 items which was used in this study. A summary of the patient instruction and the (final) item set are listed in the Appendix.

Study Population

Patients had to meet the following criteria: (1) be 12 years of age or older, (2) be living at home, and (3) have had an orthopedic or rheumatologic disorder of the lower extremity for at least 2 weeks. Consecutive eligible cases were sampled from (1) 7 private practices for physical therapy, (2) 4 outpatient clinics of hospitals (orthopedic, rehabilitation, rheumatology, and trauma departments), and (3) 7 outpatient clinics of rehabilitation centers. The sample strategy was intended to represent patients with different grades of functional limitations in rising and sitting down and aimed at recruiting equal numbers of patients for the three different settings.

The patients completed the self-administered questionnaire at home and gave additional information about age, sex, educational level, disorder, and the way the questionnaire was filled in (alone or with help from another person). The information about the disorder was checked by their doctors or therapists.

Item Scaling and Reduction

The postulated existence of 5 object dimensions (high chair, low chair, toilet, bed, and car) and 1 global dimension was tested using Mokken scale analysis,^{20,21} which can be viewed as a probabilistic version of Guttman scale analysis or more generally as a nonparametric approach to item response theory. Mokken scale analysis assumes the existence of a latent unidimensional scale (eg, "functional limitations") represented by a set of dichotomous items related to this scale. When scale criteria are met (see below), the respondents can be ordered with respect to this latent scale by means of the item set. Furthermore, the items can also be ordered hierarchically with respect to this latent scale. Thus, when items representing serious functional limitations are answered affirmatively by the respondents, items representing less serious limitations will also be predominantly answered affirmatively by the respondents.

Scale criteria are met when all coefficients of scalability for pairs of items (H_{ij}) are positive, while the scalability coefficients for the items in relation to the scale at issue (H_i) and for the whole item set (H) do not fall below a positive constant (c) chosen by the investigator. A minimum value of $c = .30$ is recommended,²¹ but higher values for H_i and H imply fewer violations and thus a better hierarchy. A rule of thumb is to speak of a "strong scale" for $H \geq .50$, of a "moderate scale" for $.40 \leq H < .50$, and of a "weak scale" when $.30 \leq H < .40$.²¹ H is based on the number of "correct" answers by the respondents to the items representing less serious limitations, given their answers to items representing serious limitations. More specifically, H equals one minus the number of correct answers expected (by answer frequency) divided by the number of correct answers found.

When these scale criteria are met, the order of the subjects on the latent scale corresponds with the number of affirmative responses the subjects give to the item set. This is called the sum score. To permit mutual comparison of scales with different numbers of items, the scale sum scores are standardized, ie, scores (range 0 to 10) are calculated as the proportion of the total possible score for the scale at issue multiplied by ten. The hierarchical order of the items corresponds with the order of the proportion of subjects responding affirmatively to the items, as expressed in the item mean score.

Robustness

Robustness is concerned with differences between subgroups of patients in scalability of the item set.^{20,22} Differences in scalability induce bias when comparing subgroups. In this study, subgroups were distinguished by (1) age (younger than 55 years versus 55 years and older), (2) sex, (3) educational level (primary versus secondary school), (4) localization of the disorder (unspecified, foot/ankle, knee, hip, multiple localizations), (5) type of disorder (unspecified, soft tissue injury, postfracture, osteoarthritis, rheumatoid arthritis, postoperative, amputation, peripheral neurologic, multiple disorders), (6) setting (private practice for physical therapy, hospital, rehabilitation center) and (7) questionnaire administration (filled in alone versus with the help of another person). Robustness was tested by performing a significance test ($\alpha < .05$) for independent H values of the subgroups²³ based on the coefficient of scalability H and its standard error.²⁰

Reliability

Reliability refers to the reproducibility of measurements with an instrument.²⁴ The intratest reliability was determined by calculating the reliability coefficient Rho . In a typical Mokken scale there is a substantial variation in the level of item difficulties. In such cases Cronbach's alpha strongly underestimates the intratest reliability.²⁰

Test-retest reliability was determined on a subgroup of patients from the rehabilitation outpatient hospital clinic. This subgroup consisted of 28 patients with stable disease, according to their doctors. These patients filled in the questionnaire twice with a 1-week interval. To estimate the test-retest reliability of the scale sum scores the intraclass correlation coefficient was calculated.²⁶

Content Validity

Content validity refers to the completeness with which an index covers the important areas of the domain that it is attempting to represent.¹⁹ The questionnaire was based on an extensive literature review of rising and sitting down, resulting in the operationalization of limitations in rising and sitting down, focusing on different objects (eg, high chair) and on different aspects (eg, velocity). Furthermore, patients and experts were invited during pretesting to check the item set for completeness. After item reduction, coverage of the operationalized objects and aspects was checked once more.

Construct Validity

Construct validity is concerned with the extent to which a particular measure relates to other measures in a manner that is consistent with theoretically derived hypotheses concerning the concepts that are measured.¹⁹ Before examining our data we formulated seven hypotheses: the functional limitation sum scores for all scales will show significant ($\alpha < .05$) positive correlations with (1) a doctor's or therapist's global assessment of functional limitations (as measured on a 10-point scale),

Table 1: Characteristics of the Study Population (n = 345)

Age	
Mean (SD)	52 (21)
Range	14-92
Gender (%)	
Female	56
Male	44
Educational level (%)	
Primary school	26
Secondary school	74
Localization of disorder (%)	
Foot	6
Ankle	6
Lower leg	5
Knee	29
Upper leg	2
Hip	12
Back	2
Multiple localizations	9
Unspecified	29
Type of disorder (%)	
Soft tissue injury	11
Postfracture	7
Osteoarthritis	10
Pain (of unknown origin)	3
Rheumatoid arthritis	13
Postoperative	18
Amputation	7
Peripheral neurologic	5
Multiple disorders	8
Unspecified	18
Setting (%)	
Private practice for physical therapy	26
Hospital	41
Rehabilitation center	33
Questionnaire administration (%)	
Filled in alone	82
Filled in with help from another person	18

(2) self-assessed functional limitations in walking outdoors and walking stairs³ (as measured by a preliminary questionnaire about functional limitations in walking and walking stairs), (3) self-assessed walking distance and time, and self-assessed number of flights of stairs walked (as measured by a 13-, 11-, and 9-point scale, respectively), and (4) the age of the patient.⁴ Furthermore, functional limitations sum scores for all scales will show no statistically significant ($\alpha < .05$) difference between (5) male and female patients. Finally, functional limitations sum scores for all scales will be significantly ($\alpha < .05$) smaller (6) for private practice patients than for hospital patients and (7) for hospital patients than for rehabilitation center patients.

Doctors and therapists were instructed to assess functional limitations independently from patients' responses to the questionnaire. To test the hypotheses, Spearman's rank correlation coefficient (one-tailed) was used, as well as the Mann-Whitney test (hypothesis 5, two-tailed; hypotheses 6 and 7, one-tailed).

RESULTS

Patients

Between 1991 and 1993 the QR&S was filled out at home by 345 patients living in the Amsterdam region. The study population characteristics are shown in table 1.

Scalability

Mokken scale analysis confirms the existence of five object dimensions. Five strong ($H \geq .53$) hierarchical unidimensional scales can be formed, with 6 (high chair), 7 (low chair), 6

(toilet), 6 (bed), and 7 (car) items, respectively (table 2). The severity of functional limitations increases in the scales from top to bottom. For example, in scale 1 (high chair) minor limitations in rising and sitting down involve the use of the arms during rising. An increase in limitation involves the use of arms during sitting down. A further increase results in a longer duration of rising and makes patients sit only in high chairs with armrests. Severe limitations bring about sitting down in a different manner (involving dropping) and a longer duration of sitting down.

Furthermore, Mokken scale analysis confirms the existence of one global dimension underlying the five object dimensions, albeit, with the characteristics of a weak scale ($H = .39$, data not shown). Therefore, to obtain a strong scale ($H \geq .50$), the less well-fitting items were removed. This resulted in a first combination scale containing 14 items about functional limitations in rising and sitting down with respect to high chairs, toilets, and beds (table 2). Mokken scale analysis of the removed items revealed a second strong combination scale. This second combination scale contains nine items about low chairs and cars (table 2). The resulting object and combination scales comprise 32 (out of 54) items.

With regard to the resulting scales, it can be observed that items about both rising and sitting down regarding a specific object fit into the same scales. This phenomenon applies to all scales. Furthermore, it can be noted from the similarly worded items (which differ only with respect to rising or sitting down) that minor limitations involve problems with rising, whereas with an increase in limitations problems with sitting down are evoked as well (eg, table 2: items 1 and 5). This pattern applies to all items. Finally, it is observed that, with respect to the objects high chair (1), low chair (2), toilet (3), and bed (4), minor limitations involve the use of arms during rising, whereas an increase in limitations results in rising in a different manner or a longer duration of rising (eg, table 2: items 2, 6, and 18, and items 4, 11, and 12). The same pattern is found for sitting down.

Robustness

Robustness of all object and combination scales was tested (table 3). Scalability does not differ significantly with respect to sex, educational level, localization of disorder, and setting. Scalability differs significantly with respect to age for two scales (bed, $p < .01$; combination scale 1, $p < .01$) and questionnaire administration for one scale (toilet, $p < .05$). Differences in scalability with respect to type of disorder (high chair, $p < .001$; low chair, $p < .01$; bed, $p < .05$) are mainly attributable to the patients with rheumatoid arthritis. The H values for this patient group are relatively low (high chair, $H = .39$; low chair, $H = .27$; bed, $H = .35$). Results indicate that the scales are in general robust.

Reliability

The reliability coefficient Rho for object and combination scales ranges from .77 to .91 (table 2), indicating that the intratest reliability of the scales is good.²⁵ The intraclass correlation coefficient ranges from .72 to .90 (table 3), indicating test-retest reliability of the scale sum scores is in general sufficient.²⁶

Content Validity

The object scales 1 to 5 comprise 32 items in total. Concerning the operationalized objects, 6 statements related to high chairs, 7 to low chairs, 6 to toilets, 6 to beds, and 7 to cars. Furthermore, with respect the operationalized aspects, 10 statements related to velocity, 13 to use of arm(rest)s, none to use of help, and 9 to "other differences" in performance.

Table 2: Scalability (According to Mokken Scale Analysis) of the 32 QR&S Items, and the Scalability, Intratest Reliability, and Median (IQR) Standardized Sum Scores of the QR&S Scales

Item*	Mean†	Scale					Combination Scale	
		1. High Chair	2. Low Chair	3. Toilet	4. Bed	5. Car	1. High Chair, Toilet, and Bed	2. Low Chair and Car
Minor Limitations‡								
1. Car, R in a different way (13)	.57					.50		
2. Bed, R always with arms (11)	.56				.62		.63	
3. Car, R longer (12)	.56					.56		.54
4. Low chair, R always with arms (5)	.54		.66					.52
5. Car, SD in a different way (27)	.53					.53		
6. Bed, R via edge (10)	.50				.60		.54	
7. High chair, R always with arms (2)	.48	.71					.57	
8. Car, SD longer (26)	.47					.55		.49
9. Car, R always with both arms (14)	.46					.50		
10. Toilet, R always with holding fast (8)	.44			.68			.56	
11. Low chair, R longer (3)	.44		.54					.50
12. Low chair, R via edge (4)	.42		.53					.47
13. Toilet, SD always with holding fast (22)	.40			.69			.58	
14. Bed, SD always with arms (25)	.37				.51		.55	
15. Low chair, SD always with arms (20)	.37		.53					.52
16. High chair, SD always with arms (17)	.36	.64					.55	
17. Low chair, SD always with dropping (19)	.34		.44					
18. Bed, R longer (9)	.33				.49			
19. Car, SD always with both arms (29)	.30					.59		.41
20. Toilet, R longer (6)	.28			.57			.51	
21. High chair, R longer (1)	.26	.56					.48	
22. Low chair, SD longer (18)	.25		.54					.55
23. Toilet, SD longer (21)	.22			.57			.53	
24. High chair, S with armrests (30)	.21	.52					.48	
25. Car, SD always with dropping (28)	.20					.53		.47
26. Bed, SD longer (24)	.17				.62		.53	
27. High chair, SD always with dropping (16)	.15	.49						
28. Toilet, R via edge (7)	.15			.49			.49	
29. High chair, SD longer (15)	.13	.57					.56	
30. Low chair, S with armrests (31)	.13		.39					
31. Toilet, S with armrests (32)	.12			.52				
32. Bed, SD "extra high" (23)	.11				.40			
Severe Limitations								
Coefficient of scalability H		.58 [†]	.53	.59	.55	.54	.54	.50
Coefficient of reliability Rho		.81**	.82	.85	.77	.83	.91	.84
Standardized sum score								
Median		1.7	2.9	1.7	3.3	4.3	2.9	3.3
IQR		0.0-5.0	0.0-5.7	0.0-5.0	0.0-5.0	1.4-7.1	0.7-5.7	1.1-6.7

Abbreviations: IQR, interquartile range; R, rising; S, sit; SD, sitting down.

* Abbreviated items. The numbers between brackets refer to the nonabbreviated items in the appendix.

† Proportion of subjects responding affirmatively to the items.

‡ Severity of functional limitations increases in a scale from top to bottom.

|| Coefficient of scalability H_i for the scalability of the items in relation to the scale (range 0 to 1). A minimum value of H_i = .30 is recommended.

† Coefficient of scalability H for the scalability of an item set (range 0 to 1). H ≥ .50 indicates a strong scale.

** Coefficient Rho for the intratest reliability of an item set. Rho ≥ .75 is minimally required.

^{||} Sum scores are standardized by dividing the sum score by the corresponding maximal sum score and multiplying it by 10 (range 0 to 10).

The combination scales 1 and 2 comprise 23 statements in total. From these, 5 relate to high chairs, 5 to low chairs, 5 to toilets, 4 to beds, and 4 to cars. Furthermore, 9 statements relate to velocity, 10 to use of arm(rest)s, none to use of help, and 4 to "other differences" in performance. These results indicate that the operationalized objects are covered well by the reduced item set, in both object and combination scales. The same applies to the operationalized aspects, with the exception of the aspect of help.

Construct Validity

Scale sum scores correlate .30 to .41 ($p < .001$) with doctor's or therapist's global assessment of functional limitations (table 3). Sum scores correlate .20 to .59 ($p < .001$) with self-assessment of functional limitations in walking outdoors and walking stairs. Scale sum scores correlate .17 ($p < .01$) to .58 ($p < .001$) with self-assessment of walking distance and time, and number of flights of stairs walked (table 3). Low chair sum score does not correlate significantly with age. The

other sum scores correlate .15 ($p < .01$) to .44 ($p < .001$) with age. All except two correlation coefficients are between .20 and .60, which are usual values when testing construct validity.²⁷

All scale sum scores do not differ significantly ($p < .05$) between male and female patients (table 3). All scale sum scores are significantly ($p < .05$) smaller for private practices patients than for hospital patients. Low chair (median standardized sum score 2.9 and 4.3, $p = .14$), car (4.3 and 5.7, $p = .11$), and combination scale 2 (3.3 and 4.4, $p = .09$) scale sum scores do not show a significant difference between hospital patients and rehabilitation center patients. The other sum scores are significantly ($p < .001$) smaller for hospital patients than for rehabilitation center patients.

Thus, results from construct validity testing were as hypothesized for 5 out of 7 hypotheses. The fourth hypothesis (age) was rejected for 1 out of 7 scales (low chair). Hypothesis 7 (hospital versus rehabilitation center) was rejected for 3 out of 7 scales (low chair, car, and combination scale 2).

Table 3: Robustness, Test-Retest Reliability, and Construct Validity of the QR&S Scales

	n ¹	Scale					Combination Scale	
		1. High Chair	2. Low Chair	3. Toilet	4. Bed	5. Car	1. High Chair, Toilet, and Bed	2. Low Chair and Car
Robustness [†]	345							
(1) Age		.54-.59	.52-.54	.52-.62	.45-.64**	.50-.57	.45-.59**	.49-.50
(2) Sex		.51-.64	.50-.57	.54-.66	.52-.58	.54-.56	.53-.55	.46-.54
(3) Educational level		.56-.64	.51-.57	.52-.63	.48-.57	.52-.55	.48-.55	.50-.50
(4) Localization of disorder		.44-.74	.40-.60	.41-.59	.39-.63	.43-.70	.46-.54	.43-.66
(5) Type of disorder		.39-.89***	.35-.80**	.49-.79	.27-.92***	.41-.67	.38-.68	.38-.59
(6) Setting		.56-.57	.53-.54	.57-.60	.48-.68*	.47-.53	.50-.56	.43-.50
(7) Questionnaire administration		.57-.57	.51-.60	.40-.62*	.45-.55	.49-.55	.42-.54	.49-.50
Test-retest reliability [‡]	28	.79 [§]	.72	.90	.76	.81	.89	.75
Construct validity	223	.38 [§]	.30	.35	.37	.35	.41	.34
(1) Doctor/therapist's global assessment	345	.56 [§]	.20	.53	.48	.35	.59	.30
(2) Self-assessment of Walking outdoors	345	.47	.21	.50	.42	.40	.52	.32
(3) Self-assessment of Walking distance	319	.54 [§]	.17	.53	.48	.37	.58	.29
(4) Age	325	.50	.26	.48	.46	.42	.53	.38
(5) Male vs female	310	.47	.22	.46	.56	.28	.53	.26
(6) Private practice vs hospital	342	.41	n.s.	.38	.33	.21	.44	.15
(7) Hospital vs rehabilitation center	149/193	1.7/1.7 [¶]	2.9/2.9	1.7/1.7	3.3/3.3	4.3/4.3	2.1/2.9	3.3/3.3
	90/140	0.0/1.7***	1.4/2.9*	0.0/1.7***	0.0/3.3***	2.9/4.3***	0.7/2.9***	2.2/3.3**
	140/115	1.7/3.3***	2.9/4.3	1.7/3.3***	3.3/5.0***	4.3/5.7	2.9/5.0***	3.3/4.4

Abbreviation: NS, not significant.

* $p < .05$, ** $p < .01$, *** $p < .001$.

¹ Number of subjects.

[†] Robustness with respect to 7 patient characteristics. The 7 patient characteristics (and distinguished groups) are (1) age (< 55 years vs ≥ 55 years), (2) sex, (3) educational level (primary vs secondary school), (4) localization of the disorder (unspecified, foot/ankle, knee, hip, multiple localizations), (5) type of disorder (unspecified, soft tissue injury, postfracture, osteoarthritis, rheumatoid arthritis, postoperative, amputation, peripheral neurologic, multiple disorders), (6) setting (private practice for physical therapy, hospital, rehabilitation center), and (7) questionnaire administration (filled in alone vs with the help of another person).

Range of the scalability coefficients H and level of significance of test for independent H values.

[‡] Test-retest reliability of scale sum scores.

[§] Intraclass correlation coefficient.

^{||} Construct validity as determined by testing seven hypotheses: the scale sum scores will show significant positive correlations with (1) a doctor's or therapist's global assessment of limitations, (2) self-assessed limitations in walking outdoors and walking stairs, (3) self-assessed walking distance and time, and self-assessed number of flights of stairs walked, and (4) age. Furthermore, scale sum scores will show no significant difference between (5) male and female patients. Finally, scale sum scores will be significantly smaller (6) for private practice patients than for hospital patients and (7) for hospital patients than for rehabilitation center patients.

[¶] Spearman's rank correlation coefficient ($p < .05$).

^{¶¶} Median standardized sum scores and level of significance of the Mann-Whitney test.

DISCUSSION

The QR&S was intended to be a discriminative index. First, such an index requires a carefully selected item pool. Items were carefully selected by distinguishing and operationalizing different objects and aspects about rising and sitting down. Second, response sets must facilitate uniform interpretation. Therefore, dichotomous response options were chosen. A third requirement is that redundant items are deleted. According to Kirshner and Guyatt,¹⁹ ensuring that the instrument meets cumulative scaling criteria is a very powerful method to reduce the number of items. Therefore the item set from the QR&S was reduced using Mokken scale analysis. Fourth, a discriminative index requires a large and stable intersubject variation. This was demonstrated for our item set through a sufficient test-retest reliability of the scale sum scores. Further research will have to confirm the stability of the intersubject variation because the 1-week interval for testing of test-retest reliability was relatively short. Fifth, a discriminative index requires cross-sectional construct validity. This was made plausible for our item set because results from hypotheses testing were in general as hypothesized.

Furthermore, the QR&S had to provide both a comprehensive and detailed, as well as a more global picture of perceived and actual functional limitations in rising and sitting down. To create a comprehensive and detailed measure five objects dimensions were assumed to be present. The existence of these

five object dimensions was confirmed by the results of the scale analyses. The resulting object scales (high chair, low chair, toilet, bed, and car) fulfilled the criteria for strong hierarchical scales. To create a global measure the existence of one global dimension was assumed. Although our assumption concerning one global dimension was not rejected, we were not satisfied with the resulting weak scale. Therefore, two strong hierarchical "combination scales" were distinguished. Combination scale 1 (high chair, toilet, and bed) and combination scale 2 (low chair and car) are easy to interpret because they contain different objects. Further research will have to confirm the existence of two combination scales. Furthermore, which factors make the item set fall apart into two combination scales will have to be investigated. Since functional limitations seem to be the result of impairment, intraindividual, and extra-individual factors,²⁰ the existence of two combination scales can be caused by each of these factors. Impairment factors, such as differences in range of motion of the major joints among patients, might cause the existence of two combination scales, whereby combination scale 1 contains the high, moderate range of motion-requiring activities and combination scale 2 contains low, large range of motion-demanding activities. Similarly, intraindividual factors, such as differences among patients with respect to low chair and car avoidance (related to fear for accidental falls), may cause the existence of two combination scales. In this interpretation, combination scale

1 contains hard-to-avoid activities and combination scale 2 easy-to-avoid activities.

The resulting object and combination scales provide an insight into the way behavioral changes relate to grade of functional limitation. Patients had more problems with rising than with sitting down, which confirms the study of Munton.² The finding that minor limitations involve the use of arms during sitting down while an increase of limitations results in sitting in a different manner is in agreement with the item about sitting down in Tinetti's performance test.¹⁵

Moreover, to create a comprehensive and detailed measure, different aspects of rising and sitting down were operationalized. After item reduction, coverage of the operationalized aspects by the reduced item set was still considered to be sufficient. The aspect of use of help was not covered by the reduced item set. The items about use of help were dropped because of their low mean scores (in general below 3%), which made them inefficient in discriminating patients with functional limitations. These low mean scores can be attributed partially to the use of Mokken scale analysis, which requires extremely formulated items (eg, *I always get help*, etc).

Finally, the QR&S had to be applicable to outpatients with different grades of functional limitations and different types of orthopedic or rheumatological lower-extremity disorders. The study population sampled from different settings did actually represent patients with different grades of functional limitations, as was illustrated during construct validity testing. Moreover, scales proved to be robust with respect to setting. Furthermore, the study population consisted of different groups with respect to localization and type of disorder. The instrument proved to be fairly robust with respect to the different localizations and types of disorders, suggesting that the scales can be generalized to similar patient groups. Problems with respect to robustness were mainly restricted to the patients with rheumatoid arthritis. For this patient group the object scales 1, 2, and 5 proved to be less homogeneous. Therefore, for the patients with rheumatoid arthritis we recommend the use of the combination scales only.

In summary, the QR&S fulfils the requirements for a discriminative index and appears capable of distinguishing patients with varying degrees of limitations in rising and sitting down. The instrument is applicable to outpatients with different grades of functional limitations and different types of orthopedic or rheumatological lower-extremity disorders, making it suitable for comparisons across different diagnostic categories. For a detailed assessment (eg, clinical studies) the use of the object scales is recommended. The use of the combination scales is recommended when only a global picture of functional limitations is needed (eg, in survey studies).

Acknowledgment: For their commentary on the draft version of the questionnaire or for their contribution to the data collection we would like to thank (in alphabetic order): Ms. W. Bannink, A.F. de Bruin, Ms. A. van Dassel-Larooy, P.D. Deutz, A.J. van Dijk, J.W. van der Eijken, M.K. Galenkamp, T. van Gils, N. Groot, K.J. Hamelynck, B. Heeling, Ms. I. Hermans, M.A. Hoelen, H.J.J. Jöris, T. van Klaveren, P. Kühlmann, H. Laman, Ms. A. Linssen, B. Nobels, Ms. G. Schutte, W. Tolman, Ms. N. Vijn, Ms. T. Voogelaar, Ms. M. Winter, L. de Witte, W. Yspeert, and R. Zondervan. We also thank L. de Witte for his commentary on a draft version of this manuscript.

References

- Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991; 39:142-8.
- Munton JS, Ellis MI, Chamberlain MA, Wright V. An investigation into the problems of easy chairs used by the arthritic and the elderly. *Rheumatol Rehabil* 1981;20:164-73.
- Netherlands Central Bureau of Statistics. Physical disability in the population of the Netherlands 1986/1988. The Hague: SDU Publishers/CBS Publications, 1990.
- Lee P, Jasani MK, Dick WC, Buchanan WW. Evaluation of a functional index in rheumatoid arthritis. *Scand J Rheumatol* 1973; 2:71-7.
- Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980;23:137-45.
- Meenan RF, Gertman PM, Mason JH. Measuring health status in arthritis. The arthritis impact measurement scales. *Arthritis Rheum* 1980;23:146-52.
- Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2. The content and properties of a revised and expanded Arthritis Impact Measurement Scales Health Status Questionnaire. *Arthritis Rheum* 1992;35:1-10.
- Aichroth P, Freeman MA, Smillie IS, Souter WA. A knee function assessment chart. From the British Orthopaedic Association Research Sub-Committee. *J Bone Joint Surg Br* 1978;60-B:308-9.
- Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988;15:1833-40.
- Budiman-Mak E, Conrad KJ, Roach KE. The Foot Function Index: a measure of foot pain and disability. *J Clin Epidemiol* 1991;44: 561-70.
- Johanson NA, Charlson ME, Szatrowski TP, Ranawat CS. A self-administered hip-rating questionnaire for the assessment of outcome after total hip replacement. *J Bone Joint Surg Am* 1992;74:587-97.
- Jefferys M, Millard JB, Hyman M, Warren MD. A set of tests for measuring motor impairment in prevalence studies. *J Chronic Dis* 1969;22:303-19.
- Jebsen RH, Trieschmann RB, Mikulic MA, Hartley RB, McMillan JA, Snook ME. Measurement of time in a standardized test of patient mobility. *Arch Phys Med Rehabil* 1970;51:170-5.
- Imms FJ, Edholm OG. The assessment of gait and mobility in the elderly. *Age Ageing* 1979;8:261-7.
- Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. *J Am Geriatr Soc* 1986;34:119-26.
- Minaire P, Flores JL, Cherpin J, Weber D. The functional aptitude of a natural sample of population. A new approach to handicap (Saint Cyr-sur-le-Rhone, September 1984). *Int Disabil Stud* 1987; 9:65-8.
- Gerety MB, Mulrow CD, Tuley MR, Hazuda HP, Lichtenstein MJ, Bohannon R, et al. Development and validation of a physical performance instrument for the functionally impaired elderly: the Physical Disability Index (PDI). *J Gerontol* 1993;48:M33-8.
- Smith DS, Clark MS. Competence and performance in activities of daily living of patients following rehabilitation from stroke. *Disabil Rehabil* 1995;17:15-23.
- Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chron Dis* 1985;38:27-36.
- Mokken RJ. A Theory and Procedure of Scale Analyses. The Hague: Mouton, 1971.
- Debets P, Brouwer E. User's manual MSP: a program for Mokken scale analysis for polychotomous items; version 1.50. Groningen: iec Programma, 1989.
- Niemöller K, van Schuur W. Stochastic models for unidimensional scaling: Mokken and Rasch. In: McKay D, Schofield N, Whitely P, editors. Data analysis and the social sciences. London: Frances Pinter, 1983:120-70.
- Marascuilo LA. Large-sample multiple comparisons. *Psychol Bull* 1966;65:280-90.
- Fletcher RH, Fletcher SW, Wagner EH. Clinical epidemiology: the essentials. Baltimore: Williams & Wilkins, 1988.
- Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press, 1989.
- Hinderer SR, Hinderer KA. Quantitative methods of evaluation. In: DeLisa JA, editor. Rehabilitation medicine: principles and practice. 2nd ed. Philadelphia: J.B. Lippincott Company, 1993:96-121.
- Bellamy N. Musculoskeletal clinical metrology. Dordrecht: Kluwer Academic Publishers, 1993:25-9.

28. Verbrugge LM, Jette AM. The disablement process. *Soc Sci Med* 1994;38:114.

**APPENDIX: STATEMENTS OF THE QR&S
AND A SUMMARY OF THE INSTRUCTIONS
FOR THE PATIENTS***

Please answer every statement with YES that (1) applies to your current situation and also (2) is connected with your health.

1. It takes me longer to get up from a chair with a high seat, eg, from a dining chair, a kitchen chair, or an office chair.
2. I always use my arms to get up from a chair with a high seat, eg, I pull myself up from the table, I push myself off of the armrests, I push myself off the seat.
3. It takes me longer to get up from a low chair or sofa, eg, from an easy chair or a deep sofa.
4. I always have to shift forward a little before I get up from a low chair or sofa.
5. I always use my arms to get out of a low chair or sofa, eg, I pull myself up from the table, I push myself off of the armrests, I push myself off the seat.
6. It takes me longer to get up from the lavatory.
7. I always shift forward a little before I get up from the lavatory.
8. I always grasp for support to get up from the lavatory, eg, the door post, the wash-basin, a handle, or an assist bar.
9. It takes me longer to get up from my bed.
10. I always shift to the edge of the bed before I get up.
11. I always use my arms to get up from the bed, eg, I hold on to something or I push myself from the bed.
12. It takes me longer to get out of a car.
13. When I get out of a car I do it in a different way, eg, I put both my legs on the ground and then I stand up.

14. I always use both hands to hold on to something while I get out of a car.
15. It takes me longer to get on a chair with a high seat, eg, on a dining chair, a kitchen chair or an office chair.
16. When I sit down onto a chair with a high seat I always let myself drop the last bit.
17. I always use my arms to get on a chair with a high seat, eg, I hold the table, I lean on the armrests, or I lean on the seat.
18. It takes me longer to sit down on a low chair or sofa, eg, on an easy chair or a deep sofa.
19. When I sit down into a low chair or sofa I always let myself drop the last bit.
20. I always use my arms to sit down on a low chair or sofa, eg, I hold the table, I lean on the armrests, or I lean on the seat.
21. It takes me longer to sit down on the lavatory.
22. I always hold on to something when I sit down on the lavatory, eg, the door post, the wash-basin, a handle, or an assist bar.
23. I only sit down on an "extra high" bed, and never on an ordinary bed.
24. It takes me longer to sit down on the bed.
25. I always use my arms to sit down on the bed, eg, I grasp hold of something, or I lean with my hands on the bed.
26. It takes me longer to get into a car.
27. When I get into a car I do it in a different way, eg, I first sit down and then I pull my legs inside.
28. When I get into a car I always let myself drop the last bit.
29. I always use two hands to hold on to something to get into a car.
30. I only sit on a chair with a high seat that has armrests and never on one without armrests.
31. I only sit on a low chair or sofa that has armrests and never on a low chair or sofa without armrests.
32. I only make use of lavatories that have assist bars and I never use lavatories that do not have assist bars.

(*Translation of the original Dutch instruction and statements. A copy of the QR&S with its manual is available on request from the first author.)